

PAT-NO: JP02002207725A

DOCUMENT-IDENTIFIER: JP 2002207725 A

TITLE: PROCESSOR FOR CONCEALING ELECTRONIC DOCUMENT

PUBN-DATE: July 26, 2002

INVENTOR-INFORMATION:

NAME	COUNTRY
FUKAZAWA, HIDEMICHI	N/A
FURUYA, HISAYOSHI	N/A

ASSIGNEE-INFORMATION:

NAME	COUNTRY
PLANET COMPUTER:KK	N/A

APPL-NO: JP2001002304

APPL-DATE: January 10, 2001

INT-CL (IPC): G06F017/21

ABSTRACT:

PROBLEM TO BE SOLVED: To conceal a specified portion on a PDF electronic document.

SOLUTION: A PDF file containing a character code is input and stored as an original copy data D1 into a storage part 20 from an original copy data inputting part 10. An original copy electronic document is displayed by a document display part 50, and a concealing area is set in one portion of the electronic document by a concealing area setting part 30. The concealing area is stored in a storage part 40 as a mask data D2. An intermediate data preparing part 60 bit-map-develops the character code in the original copy data D1, and covers its one portion by the mask data D2 to prepare an intermediate data D3 of TIFF form. The intermediate data D3 is returned again to a PDF form by a format converting part 70 to store a processing-finished data in a storage part 80. The shielding processing is executed for the PDF file output from a processing-finished data outputting part 90.

COPYRIGHT: (C) 2002, JPO

DERWENT-ACC-NO: 2002-686057

DERWENT-WEEK: 200274

COPYRIGHT 2005 DERWENT INFORMATION LTD

TITLE: Electronic document concealment processing apparatus for enterprise, converts format of intermediate data generated by performing bit map expansion of character row of data, one portion of which is concealed by mask data

PATENT-ASSIGNEE: PLANET COMPUTER KK[PLANN]

PRIORITY-DATA: 2001JP-0002304 (January 10, 2001)

PATENT-FAMILY:

PUB-NO	PUB-DATE	LANGUAGE	PAGES	MAIN-IPC
JP 2002207725 A	July 26, 2002	N/A	011	G06F 017/21

APPLICATION-DATA:

PUB-NO	APPL-DESCRIPTOR	APPL-NO	APPL-DATE
JP2002207725A	N/A	2001JP-0002304	January 10, 2001

INT-CL (IPC): G06F017/21

ABSTRACTED-PUB-NO: JP2002207725A

BASIC-ABSTRACT:

NOVELTY - A storage unit (40) stores the set up concealment area of an original document as a mask data (D2). An information generator (60) generates an intermediate data (D3) by performing bit map expansion of a character row in the stored original data (D1), one portion of which is covered by the mask data. The format of the intermediate data is converted and output as final data (D4).

USE - Electronic document concealment processing apparatus used in enterprises such as government and municipal office.

ADVANTAGE - A specific portion of the electronic document is concealed efficiently.

DESCRIPTION OF DRAWING(S) - The figure shows the block diagram of the electronic document concealment processing apparatus. (Drawing includes non-English language text).

Storage unit 40

Information generator 60

CHOSEN-DRAWING: Dwg.1/10

TITLE-TERMS: ELECTRONIC DOCUMENT CONCEAL PROCESS APPARATUS CONVERT FORMAT
INTERMEDIATE DATA GENERATE PERFORMANCE BIT MAP EXPAND CHARACTER ROW
DATA ONE PORTION CONCEAL MASK DATA

DERWENT-CLASS: T01

EPI-CODES: T01-J11A;

SECONDARY-ACC-NO:

Non-CPI Secondary Accession Numbers: N2002-541642

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開2002-207725

(P2002-207725A)

(43)公開日 平成14年7月26日(2002.7.26)

(51)IntCl'

G 06 F 17/21

識別記号

570

F I

G 06 F 17/21

テ-マコ-ト(参考)

570 M 5B009

審査請求 未請求 請求項の数7 OL (全11頁)

(21)出願番号 特願2001-2304(P2001-2304)

(22)出願日 平成13年1月10日(2001.1.10)

(71)出願人 593142927

株式会社プラネットコンピュータ
静岡県三島市芙蓉台一丁目6番22号

(72)発明者 深澤 秀通

静岡県三島市大社町5番4号 株式会社プラネットコンピュータ内

(72)発明者 古谷 寿佳

静岡県三島市大社町5番4号 株式会社プラネットコンピュータ内

(74)代理人 100091476

弁理士 志村 浩

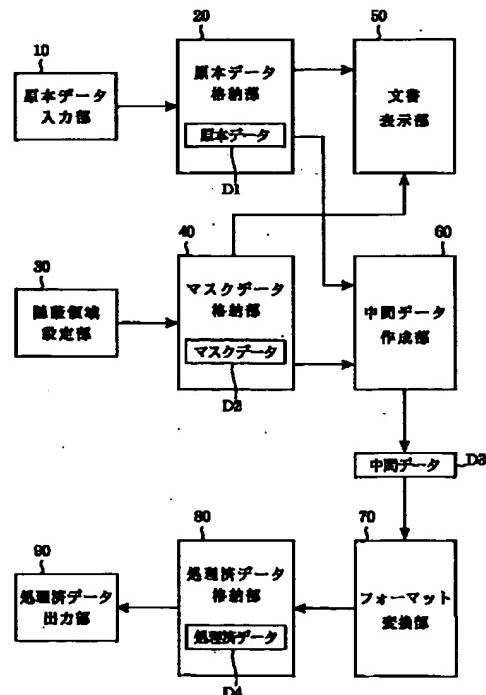
Fターム(参考) 5B009 NA04 NC06 RB34

(54)【発明の名称】電子文書に対する隠蔽処理装置

(57)【要約】

【課題】PDF電子文書に対して、特定部分を隠蔽する処理を行う。

【解決手段】原本データ入力部10から、文字コードを含むPDFファイルを、原本データD1として入力し格納部20に格納する。文書表示部50により、この原本電子文書を表示させ、隠蔽領域設定部30によって、原本電子文書上的一部分に隠蔽領域を設定する。隠蔽領域はマスクデータD2として格納部40に格納される。中間データ作成部60は、原本データD1内の文字コードをビットマップ展開し、その一部をマスクデータD2で覆うことにより、TIFF形式の中間データD3を作成する。フォーマット変換部70によって、中間データD3を再びPDF形式に戻し、処理済データD4を格納部80に得る。処理済データ出力部90から出力されるPDFファイルには、隠蔽処理が施されている。



【特許請求の範囲】

【請求項1】 所定のコードの羅列によって文字列を表す第1属性のデータと、多数の画素の集合によって画像を表す第2属性のデータと、の少なくとも2種類の属性のデータを組み合わせることにより、文字と画像とが混在した文書を定義する混在文書フォーマットを用い、この混在文書フォーマットで記述された電子文書に対して、少なくとも文字列の指定部分に対する隠蔽処理を施す隠蔽処理装置であって、

隠蔽処理の対象となる原本電子文書を、前記混在文書フォーマットで記述された原本データとして入力する原本データ入力部と、

入力した原本データを格納する原本データ格納部と、前記原本電子文書上の隠蔽すべき文字列を覆うための隠蔽領域を設定し、この隠蔽領域を示すマスクデータを生成する隠蔽領域設定部と、

生成されたマスクデータを格納するマスクデータ格納部と、

前記原本データ格納部に格納されている原本データと、前記マスクデータ格納部に格納されているマスクデータとに基づいて、ディスプレイ画面上に、前記原本電子文書と前記隠蔽領域とを重ねて表示する文書表示部と、前記原本データ格納部に格納されている原本データと、前記マスクデータ格納部に格納されているマスクデータを取り込み、原本データ内の前記第1属性のデータについては、これを前記第2属性のデータに変換することにより、前記原本電子文書に対して前記隠蔽領域部分を隠蔽した状態を示す前記第2属性のデータを作成し、前記第2属性のデータのみからなる中間文書フォーマットで記述された中間データを作成する中間データ作成部と、

前記中間文書フォーマットで記述された中間データを、前記混在文書フォーマットで記述された処理済データに変換するフォーマット変換部と、

変換された処理済データを格納する処理済データ格納部と、

この処理済データ格納部に格納されている処理済データを出力する処理済データ出力部と、
を備えることを特徴とする電子文書に対する隠蔽処理装置。

【請求項2】 請求項1に記載の隠蔽処理装置において、第2属性のデータが、個々の画素をデータビットに対応させることにより画像を表現する形式のデータ、もしくは前記形式のデータに対して所定のアルゴリズムで圧縮処理を施すことにより得られるデータ、によって構成されることを特徴とする電子文書に対する隠蔽処理装置。

【請求項3】 請求項1または2に記載の隠蔽処理装置において、

隠蔽領域設定部が、原本データを二次元平面上に展開す

る原本レイヤーと、マスクデータを二次元平面上に展開するマスクレイヤーと、を定義し、前記マスクレイヤー上に展開される図形データとしてマスクデータを生成する機能を有し、

文書表示部が、前記原本レイヤー上に展開された情報と前記マスクレイヤー上に展開された情報とを重ねて表示させる機能を有し、

中間データ作成部が、前記原本レイヤー上に展開された情報と前記マスクレイヤー上に展開された情報について、前記マスクレイヤー上に展開された情報が優先表示されるような合成演算を行うことにより中間データの作成を行う機能を有するようにしたことを特徴とする電子文書に対する隠蔽処理装置。

【請求項4】 請求項3に記載の隠蔽処理装置において、

マスクレイヤー上に、隠蔽領域を示すマスクデータとともに、当該隠蔽領域に関する注釈を記述したコメントデータを展開できるようにし、

隠蔽領域設定部に、マスクデータとコメントデータとの双方を生成する機能をもたせ、

文字表示部に、マスクデータの内容とともにコメントデータの内容をディスプレイ画面上に表示させる機能をもたせ、

中間データ作成部に、コメントデータを除外した合成演算を行うことにより、コメントデータの内容を含まない中間データの作成を行う機能をもたせたことを特徴とする電子文書に対する隠蔽処理装置。

【請求項5】 請求項1～4のいずれかに記載の隠蔽処理装置において、

「電子文書を、混在文書フォーマットで記述された原本データとして原本レイヤー上に取り込み、前記電子文書に対する注釈情報を、前記原本レイヤーとは別のアノテーションレイヤー上に付加する機能をもった編集装置」を利用して隠蔽領域設定部を構成し、この編集装置における前記アノテーションレイヤーを、マスクレイヤーとして利用することを特徴とする電子文書に対する隠蔽処理装置。

【請求項6】 請求項1～5のいずれかに記載の隠蔽処理装置において、

隠蔽領域設定部が、隠蔽領域の輪郭を示す図形データをマスクデータとして生成する機能を有し、中間データ作成部が、原本電子文書の前記隠蔽領域内部に含まれていた情報を一切含まず、かつ、前記隠蔽領域の輪郭を示す情報を含む中間データを作成する機能を有することを特徴とする電子文書に対する隠蔽処理装置。

【請求項7】 請求項1～6のいずれかに記載の隠蔽処理装置において、

混在文書フォーマットとしてPDF形式のフォーマットを行い、中間文書フォーマットとしてTIFF形式のフォーマットを用いることを特徴とする電子文書に対する

隱蔽処理装置。

【発明の詳細な説明】

[0001]

【発明の属する技術分野】本発明は電子文書に対する隠蔽処理装置に関し、特に、官公庁が公的な電子文書を公開する場合、あるいは、一般企業が業務上の電子文書を公開する場合に、公開の対象外とする特定の文字列部分に対する電子的な隠蔽処理を行う装置に関する。

[0002]

【従来の技術】コンピュータを利用した情報通信技術の発達により、官公庁や一般企業では、文書の電子化が進められている。電子文書は、磁気ディスクや光ディスクなどの電子情報記録媒体に記録して保存することができ、必要に応じて、編集、複製、プリントなどの処理をコンピュータで行うことができ、しかも、インターネットなどの通信媒体を介して配布することも可能である。このように多くのメリットを有する電子文書は、今後、あらゆる分野へと普及し、これまでの紙媒体の文書が徐々に電子文書に置き換えられてゆくものと予想される。

【0003】現在、インターネットのWebページとして提示する電子文書としては、HTML (Hyper Text Markup Language) 形式のファイルが一般化しているが、一般的なビジネスで利用される電子文書としては、PDF (Portable Document Format) 形式のファイルが最も普及している。PDF形式のファイルは、文字コードとして表現されたデータ、ビットマップイメージとして表現されたデータ、ベクトル図形として表現されたデータ、などを混在させて取り扱うことができ、しかも、異なるOSアラットホームで利用した場合にも、ほぼ共通した表示形態が得られるという利点を有するため、現在のところ、PDF形式は、ビジネスで利用される電子文書の事実上の標準フォーマットとなっている。一般に、PDF電子文書の作成、編集、閲覧などの作業を行うためには、米国アドビ社から提供されているソフトウェア「Adobe Acrobat」が広く利用されている。ワードプロセッサによって入力された文字データ、スキャナやデジカメから取り込まれた画像データ、作図ソフトウェアで作成された図形データなどの様々な文書データが、「Adobe Acrobat」を利用してPDF形式の電子文書に変換された上で、種々の媒体に記録されて配布されたり、インターネットを介して配信されたりしている。

[0004]

【発明が解決しようとする課題】上述したように、今後、官公庁や一般企業によって公開されたり、配布されたりする多くの文書が、紙媒体の文書から電子文書に置き換えられてゆくものと予想されるが、文書によっては、必ずしもすべての内容を公開できるとは限らない。たとえば、官公庁によって公開される公文書の中には、法令などの要請により、たとえば個人名など、一般には公開してはならない情報が含まれていることも少なくなく

い。同様に、一般企業が公開する文書の中にも、企業秘密に属する情報や、新製品の詳細情報など、伏せておくべき情報が含まれている場合もある。このように非公開とすべき情報が文書内に含まれていた場合、従来の紙媒体の文書であれば、該当部分に対して手作業でいわゆる「墨塗り処理」を施し、部分的に判読不能の状態にして公開するという措置が施されていた。

【0005】しかしながら、電子文書の場合、このような手作業により「墨塗り処理」を施することはできない。

10 もちろん、上述した「Adobe Acrobat」など、電子文書編集用のソフトウェアには、通常、複数のレイヤーを取り扱う機能が備わっているので、「墨塗り処理」用のための専用レイヤーを定義し、この専用レイヤー上で黒塗り処理を行えば、一見したところ、非公開とすべき特定部分を隠蔽したような処理を施すことは可能である。しかし、このような疑似的な隠蔽処理を施しただけでは、非公開とすべき部分の情報は、データとしては依然として電子文書内に残っているため、配布した電子文書内のデータが解析されると、非公開部分の内容を判読されてしまうおそれがある。

20

【0006】そこで本発明は、電子文書内の特定部分を完全に隠蔽する処理を行うことが可能な電子文書に対する隠蔽処理装置を提供することを目的とする。

[0007]

【課題を解決するための手段】(1) 本発明の第1の態様は、所定のコードの羅列によって文字列を表す第1属性のデータと、多数の画素の集合によって画像を表す第2属性のデータと、の少なくとも2種類の属性のデータを組み合わせることにより、文字と画像とが混在した文書を定義する混在文書フォーマットを用い、この混在文書フォーマットで記述された電子文書に対して、少なくとも文字列の指定部分に対する隠蔽処理を施す隠蔽処理装置において、隠蔽処理の対象となる原本電子文書を、混在文書フォーマットで記述された原本データとして入力する原本データ入力部と、入力した原本データを格納する原本データ格納部と、原本電子文書上の隠蔽すべき文字列を覆うための隠蔽領域を設定し、この隠蔽領域を示すマスクデータを生成する隠蔽領域設定部と、生成されたマスクデータを格納するマスクデータ格納部と、原

40 本データ格納部に格納されている原本データと、マスクデータ格納部に格納されているマスクデータに基づいて、ディスプレイ画面上に、原本電子文書と隠蔽領域とを重ねて表示する文書表示部と、原本データ格納部に格納されている原本データと、マスクデータ格納部に格納されているマスクデータを取り込み、原本データ内の第1属性のデータについては、これを第2属性のデータに変換することにより、原本電子文書に対して隠蔽領域部分を隠蔽した状態を示す第2属性のデータを作成し、第2属性のデータのみからなる中間文書フォーマットで

50 記述された中間データを作成する中間データ作成部と、

中間文書フォーマットで記述された中間データを、混在文書フォーマットで記述された処理済データに変換するフォーマット変換部と、変換された処理済データを格納する処理済データ格納部と、この処理済データ格納部に格納されている処理済データを出力する処理済データ出力部と、を設けるようにしたものである。

【0008】(2) 本発明の第2の態様は、上述の第1の態様に係る隠蔽処理装置において、第2属性のデータを、個々の画素をデータビットに対応させることにより画像を表現する形式のデータ、もしくはこのような形式のデータに対して所定のアルゴリズムで圧縮処理を施すことにより得られるデータ、によって構成するようにしたものである。

【0009】(3) 本発明の第3の態様は、上述の第1または第2の態様に係る隠蔽処理装置において、隠蔽領域設定部が、原本データを二次元平面上に展開する原本レイヤーと、マスクデータを二次元平面上に展開するマスクレイヤーと、を定義し、マスクレイヤー上に展開される図形データとしてマスクデータを生成する機能を有し、文書表示部が、原本レイヤー上に展開された情報とマスクレイヤー上に展開された情報を重ねて表示させる機能を有し、中間データ作成部が、原本レイヤー上に展開された情報とマスクレイヤー上に展開された情報について、マスクレイヤー上に展開された情報が優先表示されるような合成演算を行うことにより中間データの作成を行う機能を有するようにしたものである。

【0010】(4) 本発明の第4の態様は、上述の第3の態様に係る隠蔽処理装置において、マスクレイヤー上に、隠蔽領域を示すマスクデータとともに、当該隠蔽領域に関する注釈を記述したコメントデータを展開できるようにし、隠蔽領域設定部に、マスクデータとコメントデータとの双方を生成する機能をもたせ、文字表示部に、マスクデータの内容とともにコメントデータの内容をディスプレイ画面上に表示させる機能をもたせ、中間データ作成部に、コメントデータを除外した合成演算を行うことにより、コメントデータの内容を含まない中間データの作成を行う機能をもたせるようにしたものである。

【0011】(5) 本発明の第5の態様は、上述の第1～第4の態様に係る隠蔽処理装置において、「電子文書を、混在文書フォーマットで記述された原本データとして原本レイヤー上に取り込み、電子文書に対する注釈情報を、原本レイヤーとは別のアノテーションレイヤー上に付加する機能をもった編集装置」を利用して隠蔽領域設定部を構成し、この編集装置におけるアノテーションレイヤーを、マスクレイヤーとして利用するようにしたものである。

【0012】(6) 本発明の第6の態様は、上述の第1～第5の態様に係る隠蔽処理装置において、隠蔽領域設定部が、隠蔽領域の輪郭を示す図形データをマスクデータとして生成する機能を有し、中間データ作成部が、原本電子文書の隠蔽領域内部に含まれていた情報を一切含まず、かつ、隠蔽領域の輪郭を示す情報を含む中間データを作成する機能を有するようにしたものである。

タとして生成する機能を有し、中間データ作成部が、原本電子文書の隠蔽領域内部に含まれていた情報を一切含まず、かつ、隠蔽領域の輪郭を示す情報を含む中間データを作成する機能を有するようにしたものである。

【0013】(7) 本発明の第7の態様は、上述の第1～第6の態様に係る隠蔽処理装置において、混在文書フォーマットとしてPDF形式のフォーマットを用い、中間文書フォーマットとしてTIFF形式のフォーマットを用いるようにしたものである。

10 【0014】

【発明の実施の形態】以下、本発明を図示する実施形態に基づいて説明する。

【0015】§1. 装置の基本構成

はじめに、本発明に係る電子文書に対する隠蔽処理装置の基本構成を説明する。図1は、本発明の基本的な実施形態に係る電子文書に対する隠蔽処理装置の構成を示すブロック図である。この隠蔽処理装置は、隠蔽処理の対象となる電子文書（ここでは、原本電子文書と呼ぶ）に対して、文字列の指定部分に対する隠蔽処理を施し、これを処理済の電子文書として出力する機能を有している。図における原本データ入力部10は、この原本電子文書を原本データD1として入力する機能を有し、入力された原本データD1は、原本データ格納部20に格納される。オペレータは、この原本電子文書内の非公開部分に隠蔽領域を設定するため、隠蔽領域設定部30に所定の指示を与えて、隠蔽領域を示すマスクデータD2を作成する作業を行う。こうして作成されたマスクデータD2は、マスクデータ格納部40に格納される。文書表示部50は、原本データ格納部20に格納されている原本データD1に基づいて、ディスプレイ画面上に原本電子文書を表示する機能を有しており、また、マスクデータ格納部40内に格納されているマスクデータD2に基づいて、ディスプレイ画面上に隠蔽領域を表示する機能を有している。したがって、ディスプレイ画面上には、原本電子文書と隠蔽領域とが重ねて表示されることになり、オペレータは、このディスプレイ画面を見ながら、隠蔽領域設定部30に所定の指示を与え、非公開部分に隠蔽領域を設定する作業を行うことができ、その結果を確認することができる。

【0016】こうして、原本データD1に対応するマスクデータD2が作成できたら、中間データ作成部60によって中間データD3が作成される。この中間データD3は、原本データD1とマスクデータD2とに基づいて作成されるデータであり、原本電子文書に対して隠蔽領域部分を隠蔽する処理を施したデータに相当する。もともと、このような隠蔽処理を施すことにより、中間データD3のフォーマットは、原本データD1とは異なるフォーマットになる。そこで、この中間データD3に対して、フォーマット変換部70においてフォーマット変換処理を施し、原本データD1と同じフォーマットをもつ

た処理済データD4を生成し、これを処理済データ格納部80に格納する。最後に、必要に応じて、この処理済データD4を処理済データ出力部90を介して外部へと出力し、磁気記録媒体や光記録媒体などの電子記録媒体に書き込んだり、インターネットなどの通信媒体を介して配信したりすることができる。

【0017】かくして、原本データ入力部10から入力された所定のフォーマットの原本データD1は、最終的には、全く同じフォーマットの処理済データD4として、処理済データ出力部90から出力されることになる。このように、原本データD1と処理済データD4とは同じフォーマットのデータであるが、処理済データD4では、原本データD1に含まれていた一部の情報が隠蔽された状態となっている。より正確に言えば、原本データD1に含まれていた隠蔽対象部分の情報は、処理済データD4内からは抹消されており、処理済データD4をどのように解析しても、隠蔽対象部分の情報が判読されることはない。

【0018】この隠蔽処理装置を用いれば、オペレータは、原本データ入力部10に対して、隠蔽処理の対象となる原本電子文書を原本データD1として入力した上で、文書表示部50によってディスプレイ画面上に表示される内容を確認しながら、隠蔽領域設定部30に対して隠蔽領域を指定する指示を与える作業を行うだけで、処理済データ出力部90から、処理済データD4を得ることができる。オペレータが隠蔽領域の設定作業さえ行えば、隠蔽領域設定部30によるマスクデータD2の作成処理、中間データ作成部60による中間データD3の作成処理、フォーマット変換部70によるフォーマット変換処理は、すべて自動で行われ、オペレータはこれらの処理に関与する必要はない。また、原本データD1は、原本データ格納部20内にそのまま格納されているので、オペレータは、この原本データD1を利用して、何度も隠蔽処理を繰り返し実行することができる。たとえば、配布先に応じて、それぞれ隠蔽箇所を変えた複数種類の電子文書を用意したい場合であれば、同一の原本データD1に対して、複数種類のマスクデータD2を作成すれば、個々のマスクデータD2を利用して複数種類の処理済データD4を得ることができるので、これら複数種類の処理済データD4をそれぞれの配布先へと配布することができる。もちろん、原本データD1とともにマスクデータD2も格納保存されているので、処理済データD4が失われた場合でも、必要な場合には、保存されていた原本データD1とマスクデータD2に基づいて、処理済データD4を何回でも繰り返して作成することができる。

【0019】なお、図1では、説明の便宜上、この隠蔽処理装置を機能ブロックの集合として示すが、実際には、この隠蔽処理装置は、汎用のコンピュータに専用のアプリケーションソフトウェアを組み込むことにより実

現できる。あるいは、単独のアプリケーションソフトウェアとしてではなく、複数のアプリケーションソフトウェアの組み合わせとして実現することも可能である。更に、ライブラリプログラムやプラグインプログラムとして組み込むことも可能であるし、OSプログラムに組み込むことも可能である。したがって、実際には、原本データ格納部20、マスクデータ格納部40、処理済データ格納部80といった各格納部は、コンピュータに備わっている種々のデータ記録手段によって構成される。また、文書表示部50は、このコンピュータ用の表示処理用のハードウェアおよびソフトウェアによって構成され、原本データ入力部10や処理済データ出力部90は、このコンピュータ用のデータ入出力手段によって構成される。更に、隠蔽領域設定部30、中間データ作成部60、フォーマット変換部70といった各構成要素は、実際には、このコンピュータのハードウェアおよびソフトウェアを利用した処理機能によって実現される構成要素である。

【0020】§2. 装置の具体的な動作
20 続いて、この図1の隠蔽処理装置で取り扱われるデータのフォーマットに言及しながら、この装置の具体的な動作および各構成要素の具体的な機能を説明する。まず、この隠蔽処理装置によって隠蔽処理を施す対象となる電子文書は、所定のコードの羅列によって文字列を表す第1属性のデータと、多数の画素の集合によって画像を表す第2属性のデータと、の少なくとも2種類の属性のデータを組み合わせることにより、文字と画像とが混在した文書を定義する混在文書フォーマットで記述されていることを前提としている。このような混在文書フォーマットとしては、前述したPDF形式のフォーマットが一般に普及している。図2は、この混在文書フォーマット(PDF)で記述された電子文書ファイルのデータ構造を示す図である。図示のとおり、この混在文書フォーマットでは、第1属性のデータと第2属性のデータとが混在している。

【0021】PDF形式のフォーマットでは、文字コードとして表現されたデータ、ビットマップイメージとして表現されたデータ、ベクトル图形として表現されたデータ、などを混在させて取り扱うことができるため、通常、ワードプロセッサによって入力された文字列のデータは、文字コードとして取り扱われ、スキャナやデジカメから取り込まれたイメージのデータは、画素の集合からなる画像データとして取り扱われ、作図ソフトウェアなどで作成された图形のデータは、ベジェなどのベクトルデータとして取り扱われる。したがって、PDF形式のファイルには、文字コード、画像データ、ベクトルデータが混在することになる。ただ、これらのうち、少なくとも文字コードと画像データとを組み合わせることにより、文字と画像とが混在した文書を定義できるフォーマットで記述されている電子文書であれば、本発明に係

る隠蔽処理装置による隠蔽処理を実行することができる。

【0022】そこで、ここでは、原本データ入力部10によって入力される原本データD1が、図2に示すように、第1属性のデータ（所定のコードの羅列によって文字列を表す文字コードからなるデータ）と、第2属性のデータ（多数の画素の集合によって画像を表す画像データからなるデータ）と、によって構成されている例について、以下の説明を行うことにする。もちろん、本発明で取り扱う混在文書フォーマットとしては、この他に、第3属性のデータ（たとえば、ベジェなどのベクトルデータ）や、更に、第4、第5属性のデータを含んでいてもかまわない。また、本発明における第2属性のデータは、個々の画素をデータビットに対応させることにより画像を表現する形式のデータ（いわゆるビットマップ形式のデータ）そのものに限定されるわけではなく、このようなビットマップ形式のデータに対して所定のアルゴリズムで圧縮処理を施すことにより得られるデータであってもかまわない。

【0023】一方、中間データ作成部60によって作成される中間データD3は、図3に示すように、第2属性のデータのみからなる中間文書フォーマットで記述されたデータとなっている。別言すれば、中間データD3は、文字コードを含まない画像データのみからなる中間文書フォーマットで記述されることになる。このような中間文書フォーマットとしては、たとえば、TIF（Tag Image File Format）形式のフォーマットを用いることができる。これに対し、フォーマット変換部70による変換処理によって得られる処理済データD4は、原本データD1と同じ混在文書フォーマットのデータとなる。なお、隠蔽領域設定部30によって作成されるマスクデータD2は、どのようなフォーマットで記述されたデータであってもかまわないが、ここでは、図形を示すのに適したベクトルデータを用いた形式のフォーマットを用いている。

【0024】結局、この隠蔽処理装置を用いれば、混在文書フォーマット（たとえば、PDF）で記述された原本データD1から、中間文書フォーマット（たとえば、TIFF）で記述された中間データD3が作成され、この中間データD3のフォーマットを再び混在文書フォーマット（PDF）に戻すことにより、処理済データD4が得られることになる。本発明の意図するところは、混在文書フォーマット（PDF）で記述された原文電子文書に対して、同一のフォーマット（PDF）で記述された隠蔽処理済みの電子文書を得ることにあり、処理済データ出力部90から出力される処理済データD4は、原本電子文書を構成する原本データD1と同一のフォーマット（PDF）で記述された電子文書として配布することができる。

【0025】統一して、この隠蔽処理装置を用いた具体的

な隠蔽処理作業を説明しながら、この装置の動作を説明する。ここでは、PDF形式のフォーマットで記述された原本データD1を、原本電子文書として入力し、必要な隠蔽処理を施す場合を考えてみる。オペレータが、原本データ入力部10から原本データD1を入力すると、この原本データD1は、原本データ格納部20に格納される。実際には、この入力処理は、コンピュータに原本電子文書となるPDFファイルを読み込む処理として行われる。ここでは、この入力された原本データD1の内容が、図4に示すようなものであったとしよう。図4は、この原本データD1の内容、すなわち原本電子文書の内容を、ディスプレイ画面上に表示した状態を示すものである。図示のとおり、この原本電子文書は、新型パソコンの仕様説明書であり、文字列部分と画像部分によって構成される。ここでは、この仕様説明書の文字列部分が第1属性のデータ（文字コード）によって構成され、画像部分が第2属性のデータ（画像データ）によって構成されているものとする。本発明の特徴は、このような原本電子文書の文字列の指定部分に対する隠蔽処理

にある。

【0026】いま、このような新型パソコンの仕様説明書を、正式な新製品発表に先立って、関係業者にPDFファイルとして配布することになったとしよう。ただし、現段階では、図4に示されている仕様のうち、予定期間、搭載メモリ、内蔵HDの詳細な数値までは公開すべきではないため、これらの数値部分に対して隠蔽処理を施した上で、この仕様説明書をPDFファイルとして配布が必要が生じたものとしよう。この場合、オペレータは、隠蔽領域設定部30に対して所定の指示を与えることにより、原本電子文書上の隠蔽すべき文字列を覆うための隠蔽領域を設定する作業を行う。前述したように、図4に示すような原本データD1の内容は、文書表示部50の機能によって、ディスプレイ画面上に表示されるので、オペレータは、この画面上で隠蔽領域を設定する作業を行うことができる。たとえば、予定期間である「¥198,000」なる文字列を覆うための隠蔽領域は、ディスプレイ画面上で、図5に示すような矩形領域A1として設定することができる。このような矩形領域A1は、たとえば、マウスなどの入力機器を利用して、対角位置にある2頂点P1, P2を指定する入力をすることにより設定することができる。

【0027】こうして設定された矩形領域A1の情報は、マスクデータ格納部40内にマスクデータD2として格納されることになる。ここでは、この矩形領域A1の輪郭を示す図形データが、マスクデータD2として格納される。上述したように、マスクデータD2のフォーマットとしては、どのようなものを用いてもかまわないが、矩形の輪郭情報をからなるマスクデータであれば、たとえば、2頂点P1, P2の座標値を示すデータによって構成することができる。隠蔽領域設定部30は、この

11

ように、オペレータの指示に基づいて、隠蔽領域の輪郭を示す図形データをマスクデータD2として生成し、これをマスクデータ格納部40に格納する処理を実行する。文書表示部50は、原本データ格納部20に格納されている原本データD1と、マスクデータ格納部40に格納されているマスクデータD2とに基づいて、ディスプレイ画面上に、原本電子文書と隠蔽領域とを重ねて表示する処理を行うので、オペレータは、設定した隠蔽領域の位置および大きさをディスプレイ画面上で確認しながら作業を進めることができる。なお、図5では、矩形領域A1の輪郭を構成する外枠線のみをディスプレイ画面上に表示させた例を示したが、たとえば、図6の矩形領域A1*のように、隠蔽領域の内部にハッチングを施したり、あるいは着色して示したりしてもよい。もちろん、隠蔽領域の内部を完全に塗り潰していくわゆる「墨塗り処理」を施した状態を示すようにしてもかまわない（この場合、隠蔽処理対象となった文字列は、ディスプレイ画面上では判読できなくなる）。

【0028】図7は、このようにして、3箇所に隠蔽領域（矩形領域）A1、A2、A3を設定したときのディスプレイ画面上の表示例を示している。図示のように、ディスプレイ画面上には、設定した隠蔽領域A1、A2、A3が矩形の輪郭線として表示されるので、オペレータは、隠蔽領域設定作業の結果を直ちに確認することができる。図8は、このような隠蔽領域設定作業を行った場合に生成されるマスクデータD2の内容を示す図である。図示のとおり、生成されるマスクデータD2は、隠蔽領域A1、A2、A3に対応する矩形を示す図形データになる。この図形データは、少なくとも各隠蔽領域の輪郭に関する情報を有していればよい。

【0029】オペレータは、必要な隠蔽領域の設定が完了したら、この隠蔽処理装置に対して、隠蔽領域設定作業の終了を示す指示を与える。この指示を受けて、中間データ作成部60は、中間データD3を作成する処理を実行する。すなわち、中間データ作成部60は、原本データ格納部20に格納されている原本データD1と、マスクデータ格納部40に格納されているマスクデータD2を取り込み、原本データD1内の第1属性のデータ（文字コード）については、これを第2属性のデータ（画像データ）に変換することにより、原本電子文書に対して隠蔽領域部分を隠蔽した状態を示す第2属性のデータ（画像データ）を作成し、第2属性のデータ（画像データ）のみからなる中間文書フォーマットで記述された中間データD3を作成する処理を行う。

【0030】この中間データ作成部60による中間データ作成処理は、具体的には、次のような処理になる。まず、原本データD1内の第1属性のデータ（文字コード）については、第2属性のデータ（画像データ）に変換する処理が行われる。図4に示す例の場合、PDFファイルとして与えられた原本データD1のうち、仕様説

12

明文の部分は、第1属性のデータ、すなわち、文字コードとして組み込まれている。そこで、この文字コード部分を画像データに変換する処理が実行される。文字コードを画像データに変換する処理は、文字を画面表示する場合や、文字をプリントする場合に行われている一般的な処理であるため、ここでは詳細な説明は省略するが、予め用意されたフォントデータに基づいて、個々の文字をビットマップ形式のデータに展開することによって行われる。こうして、文字コードを画像データに変換する10と、図4に示されているすべてのオブジェクトは、画像データ、すなわち画素の集合によって表現されることになる。

【0031】続いて、この図4に示されているような原本電子文書の内容に対して、前述の作業で設定された隠蔽領域A1、A2、A3の部分を隠蔽した状態を示す第2属性のデータ（画像データ）が作成される。具体的には、図4に示すような1枚の画像データと、図8に示すような1枚の画像データ（各隠蔽領域A1、A2、A3の内部には、隠蔽領域内であることを示す画素が配置されており、背景部分の画素とは区別される。）と、の間の画像演算処理が行われ、図9に示すように、隠蔽領域A1、A2、A3の部分が隠蔽された1枚の画像データが作成される。ここに示す例では、図9に示す画像データは、TIFF形式の画像データファイルとして作成される。この図9に示す画像データが、中間データ作成部60によって作成される中間データD3である。中間データD3は、図3に示す中間文書フォーマット（TIFF）で記述されており、第2属性のデータ（画像データ）のみから構成されている。この図9に示す画像データは、1枚の画像データ（二次元平面上の画素配列からなるデータ）であり、隠蔽領域A1、A2、A3内は塗り潰された状態になっている。別言すれば、この中間データD3においては、図7に示す各隠蔽領域A1、A2、A3の内部の文字列の情報は完全に失われていることになる。このように、中間データD3には、原本電子文書の隠蔽領域内部に含まれていた情報は、一切含まれていないことになる。

【0032】続いて、フォーマット変換部70によって、この中間データD3のフォーマットを、中間文書フォーマット（TIFF）から混在文書フォーマット（PDF）に変換する処理が行われ、処理済データD4が作成される。処理済データD4と中間データD3とは、含まれているオブジェクトの内容は全く同一であるが、相互に異なるデータフォーマットで記述されたデータということになる。処理済データD4は、図10に示すように、PDF形式のフォーマットで記述されたデータになっている。ただし、その内容を見ると、すべての情報が、第2属性のデータ（画像データ）として表現されており、第1属性のデータ（文字コード）は全く含まれていない。すなわち、この処理済データD4は、PDF

ファイルの形態をとっているものの、文字列の部分も画像の部分も、いずれも画素の集合からなる画像データとして表現されることになる。

【0033】かくして、処理済データ格納部80に格納された処理済データD4を、処理済データ出力部90から出力すれば、これをPDFファイルとして配布することが可能である。このように、原本データD1も処理済データD4も、いずれもPDFファイルという点では同等であるため、両者は全く同じ形態で配布することが可能である。別言すれば、形式的には、原本データD1の配布も処理済データD4の配布も何ら変わることはない。ただ、配布を受けた者が、これを閲覧した場合、前者では、図4に示すようにすべての情報が提示されてしまうのに対し、後者では、図9に示すように隠蔽領域内の情報が隠蔽された状態で提示されることになる。しかも、処理済データD4内には、もともと隠蔽領域内に含まれていた原本電子文書内の情報は一切含まれていないため、データをどのように解析しようとも、隠蔽された情報が判読されることはない。

【0034】なお、図9に示す例では、隠蔽領域の内部を完全に塗り潰していわゆる「墨塗り処理」を施した状態にしてあるが、中間データD3を作成する際には、隠蔽領域に対して必ずしも「墨塗り処理」を施す必要はない。要するに、処理済データD4の配布を受けた者が、これを閲覧した際に、「特定の部分が隠蔽処理されている」と認識することができればよいので、隠蔽領域は、たとえば白抜きの矩形で表示するようにしてもかまわない。要するに、中間データD3には、原本電子文書の隠蔽領域内部に含まれていた情報が一切含まれないようにし、かつ、隠蔽領域の輪郭を示す情報が含まれているようすればよい。

【0035】§3. より実用的な実施形態

最後に、本発明を実施する上で、より実用的な実施形態をいくつか述べておく。上述したように、本発明に係る隠蔽処理装置では、原本データ格納部20には、図4に示すようなオブジェクトの集合からなる原本データD1が格納され、マスクデータ格納部40には、図8に示すような图形の集合からなるマスクデータD2が格納されることになる。そして、これら2種類のデータD1, D2の内容は、文書表示部50によってディスプレイ画面上に重ねて表示され、また、中間データ作成部60によって合成される。このような処理を行うには、レイヤーを定義してデータを取り扱うと便利である。すなわち、原本データD1を二次元平面上に展開するための原本レイヤーと、マスクデータD2を二次元平面上に展開するマスクレイヤーと、を定義し、原本データD1とマスクデータD2とを、それぞれ異なるレイヤー上のデータとして取り扱うようにする。したがって、隠蔽領域設定部30は、原本レイヤーとは別個に定義されたマスクレイヤー上に展開される图形データとして、マスクデータD

2を生成する処理を行うことになる。このような取り扱いをするのであれば、文書表示部50は、原本レイヤー上に展開された情報とマスクレイヤー上に展開された情報とを重ねて表示させる処理を行うようにすればよい。また、中間データ作成部60は、原本レイヤー上に展開された情報とマスクレイヤー上に展開された情報について、マスクレイヤー上に展開された情報(マスクデータD2)が優先表示されるような合成演算を行うことにより中間データD3の生成を行えばよい。

- 10 【0036】前述したように、PDF形式の電子文書の作成、編集、閲覧などの作業を行うソフトウェアとして、「Adobe Acrobat」が広く利用されているが、この「Adobe Acrobat」にも、複数のレイヤーに分けてデータを取り扱う機能が備わっている。すなわち、「Adobe Acrobat」では、作成対象あるいは編集対象となる電子文書のデータを展開するための原本レイヤーの他に、たとえば、この電子文書に対する注釈情報を展開するためのアノテーションレイヤーが用意されている。そこで、このアノテーションレイヤーを、上述したマスクレイヤーとして利用するようにすれば、本発明に係る隠蔽処理装置の大部分を、「Adobe Acrobat」の標準的な機能を利用して実現できることになる。すなわち、図1に示す構成要素のうち、原本データ入力部10、原本データ格納部20、隠蔽領域設定部30、マスクデータ格納部40、文書表示部50、処理済データ格納部80、処理済データ出力部90の各要素は、「Adobe Acrobat」をインストールした汎用のパソコンによって構成することができる。この場合、原本データD1の入力作業は、「Adobe Acrobat」のアプリケーション上で原本電子文書となる所定のPDFファイルを開く作業として実行することができ、隠蔽領域の設定作業は、「Adobe Acrobat」のアプリケーション上で、アノテーションレイヤーを利用し、所定の矩形图形を注釈情報として入力する作業として実行することができる。「Adobe Acrobat」には、アノテーションレイヤー上の注釈情報を、原本レイヤー上の電子文書情報に重ねてディスプレイ表示する機能が備わっているので、この表示機能をそのまま文書表示部50の機能として利用することができます。また、処理済データD4は、PDFファイルであるから、「Adobe Acrobat」によって直接取り扱うことが可能なファイルであり、処理済データD4を出力する作業は、「Adobe Acrobat」のアプリケーション上あるいはOS上でのファイル操作として実行することができる。
- 20 【0037】したがって、本発明に係る隠蔽処理装置を、「Adobe Acrobat」を利用して実現するのであれば、中間データ作成部60およびフォーマット変換部70として機能する専用のソフトウェア(単独のアプリケーションとして用意してもよいし、プラグインなどの補助プログラムとして用意してもよい)を、このパソコンに組み込むだけでよい。
- 30
- 40
- 50

【0038】なお、上述したアノテーションレイヤーをマスクレイヤーとして利用するのであれば、アノテーションレイヤーの本来の機能を生かして、隠蔽領域を設けた理由などの注釈を記述するコメントデータを、このアノテーションレイヤーに書き込めるようになるとよい。電子文書中の特定の文字列を隠蔽する必要がある場合、通常、何らかの理由があるはずである。たとえば、「法令により公開が制限されている個人名の部分であるため」とか、「新製品発表前における詳細な仕様を示す数値部分であるため」といった理由により、特定の文字列部分が隠蔽処理されることになる。アノテーションレイヤーをマスクレイヤーとして利用する場合、オペレータは、アノテーションレイヤー上に隠蔽領域の輪郭に相当する図形をマスクデータとして入力する作業を行うことになるので、この作業中に、アノテーションレイヤー上に上述のような理由や他の注釈を示す文字列をコメントデータとして入力することも可能である。

【0039】結局、マスクレイヤーとして用いるアノテーションレイヤー上に、隠蔽領域を示すマスクデータとともに、当該隠蔽領域を設けた理由を記述するコメントデータを展開できるようにしておき、隠蔽領域設定部30に、マスクデータとコメントデータとの双方を生成する機能をもたせ、文書表示部50に、マスクデータの内容とともにコメントデータの内容をディスプレイ画面上に表示させる機能をもたせておけば、マスクデータとともに、コメントデータを表示させることができ、これを格納することも可能になる。

【0040】ただ、このコメントデータは、隠蔽作業を行うオペレータによって利用される情報であるため、処理済データD4上には現れないようにする必要がある。そこで、中間データ作成部60には、このコメントデータを除外した合成演算を行うことにより、コメントデータの内容を含まない中間データD3を作成する機能をもたせておくようにするとよい。具体的には、アノテーションレイヤー上に入力されたオブジェクト（矩形図形やコメント文）が、マスクデータを構成するオブジェクト（矩形図形）であるのか、コメントデータを構成するオブジェクト（コメント文）であるのか、を区別できるようにしておき、コメントデータを構成するオブジェクト（コメント文）については、その内容が中間データD3に含まれないようにすればよい。実際には、たとえば、アノテーションレイヤー上に入力されたオブジェクトのうち、図形オブジェクトはマスクデータを構成するオブジェクトであり、文字列オブジェクトはコメントデータを構成するオブジェクトである、というような基準で自動的に区別できるようにしてもよいし、オペレータが入力作業を行う際に、マスクデータを構成するオブジェクトは赤色で入力し、コメントデータを構成するオブジェクトは青色で入力する、というような規則を定めておいてもよい。あるいは、アノテーションレイヤーを更に2

枚のレイヤーに分け、マスクデータとコメントデータとをそれぞれ別個のレイヤーに入れるようにしてもよい。

【0041】以上、本発明を図示するいくつかの実施形態に基づいて説明したが、本発明はこれらの実施形態に限定されるものではなく、この他にも種々の形態で実施可能である。特に、上述の実施形態では、混在文書フォーマットとしてPDF、中間文書フォーマットとしてTIFFを用いた例を示したが、本発明を実施する上で利用するフォーマットは、必ずしもこれらのフォーマットに限定されるものではない。また、上述の実施形態では、混在文書フォーマットとして、第1属性のデータ（文字コード）と第2属性のデータ（画像データ）とを混在させたフォーマットを示したが、原本データD1には、これ以外の属性データ（たとえば、ベクトルデータ）が含まれていても問題はない。たとえば、第3属性のデータが含まれていた場合には、この第3属性のデータ部分に関しては、中間データD3を作成する段階において、ひとまず取り除いておき、処理済データD4が得られた段階において、この処理済データD4に第3属性のデータを加えるようにすればよい。

【0042】

【発明の効果】以上のとおり、本発明に係る電子文書に対する隠蔽処理装置によれば、電子文書内の特定部分を完全に隠蔽する処理を行うことが可能になる。

【図面の簡単な説明】

【図1】本発明の基本的な実施形態に係る電子文書に対する隠蔽処理装置の構成を示すブロック図である。

【図2】図1の原本データD1を記述するための混在文書フォーマット（PDF）のデータ構造を示す図である。

【図3】図1の中間データD3を記述するための中間文書フォーマット（TIFF）のデータ構造を示す図である。

【図4】図1の原本データD1の内容をディスプレイ画面上に表示した状態を示す図である。

【図5】図4に示すディスプレイ画面上において、隠蔽領域A1を設定する作業を示す図である。

【図6】図5に示す隠蔽領域A1の代わりに、内部にハッチングが施された隠蔽領域A1*を設定した状態を示す図である。

【図7】図4に示すディスプレイ画面上において、3つの隠蔽領域A1, A2, A3を設定した状態を示す図である。

【図8】図7に示す隠蔽領域A1, A2, A3の設定作業により作成されたマスクデータD2上の図形を示す図である。

【図9】図1の中間データD3の内容をディスプレイ画面上に表示した状態を示す図である。

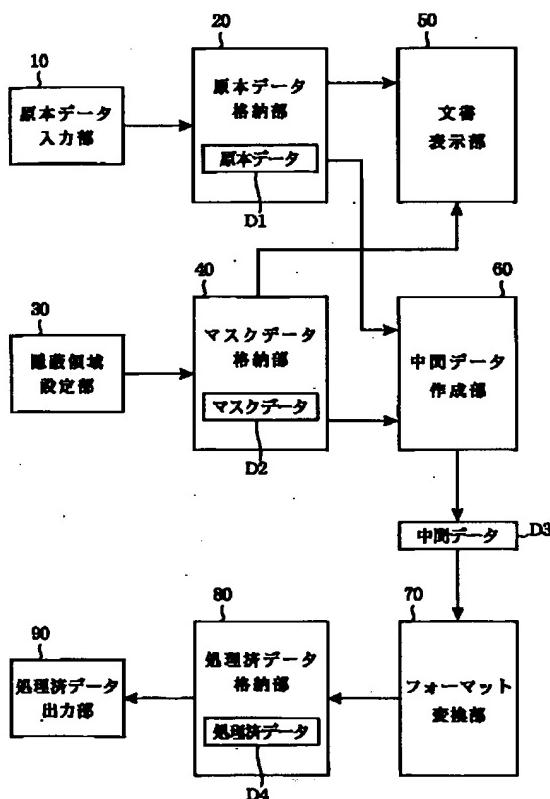
【図10】図1の処理済データD4の文書フォーマット（PDF）のデータ構造を示す図である。

【符号の説明】

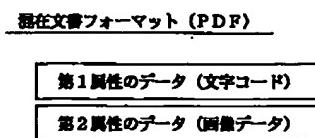
10…原本データ入力部
20…原本データ格納部
30…隠蔽領域設定部
40…マスクデータ格納部
50…文書表示部
60…中間データ作成部
70…フォーマット変換部

80…処理済データ格納部
90…処理済データ出力部
A1, A2, A3, A1*…矩形領域／隠蔽領域
D1…原本データ
D2…マスクデータ
D3…中間データ
D4…処理済データ
P1, P2…矩形領域の対角上の2頂点

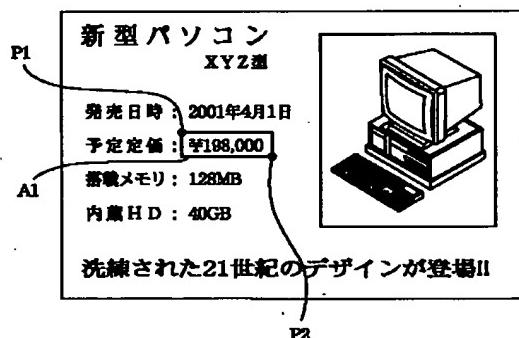
【図1】



【図2】



【図5】

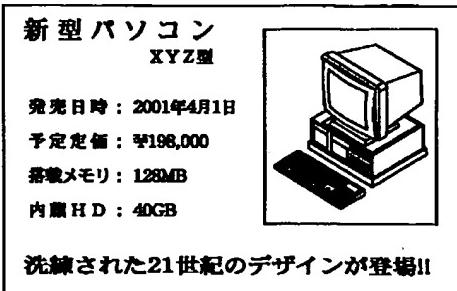


【図3】

中间文書フォーマット (TIFF)

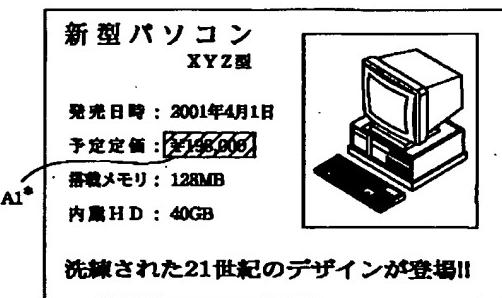
第2属性のデータ (画像データ)

D1: 原本データ (PDF)

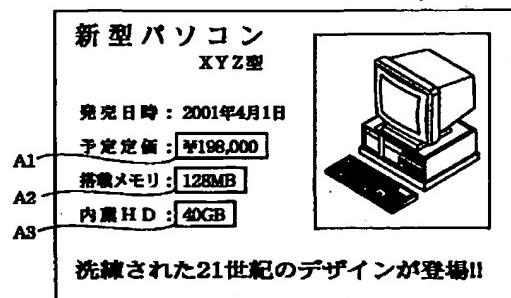


【図4】

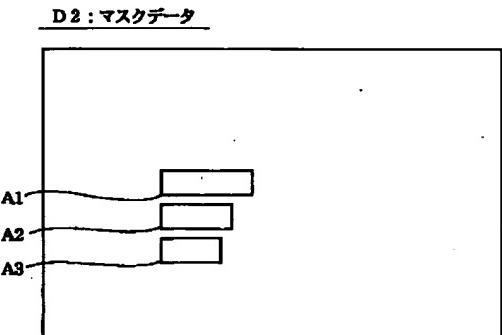
【図6】



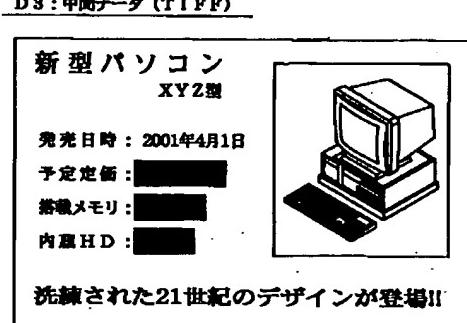
【図7】



【図8】



【図9】



【図10】

